

Weekly Report

Lu Junhua

2016 年 11 月 14 日

Done

- Went to Yvquan to talk about the netease data they possess.
- Keep collecting materials about anomalies in games (e.g., 外挂, online-game related mental disorders).
- Talked to Prof. Peng, about our next project. He gave us constructive suggestions. From the perspective of scalability or generalizability, he hoped that our next work would address a more macroscopic problem (可以上升到一定高度的, 不仅限于一个小问题的; 或者在一定程度上是完备的, 可以直接类比到现实中的问题). I will made some survey based on interviews with pals who play games with many many experiences.
- Reviewed a paper, and revised a paper.

To do

- Interview with those with game experiences and netease employees.
- Discuss about netease project with other scholars.
- Continue collecting materials (外挂game bot, mental health, bully network).

论文

- WWW2016 *Mining Online Social Data for Detecting Social Network Mental Disorders* 本文重点在于用数据挖掘方法从大量用户在互联网社交网络数据找出有(现实生活中)社交精神问题()的人. 一般来说, 找有精神问题的人, 是通过心理医生给做一个问卷来判断的, 很显然这个方法现在不实用. 作者们与心理专家紧密合作, 将传统的心理量表中的特征数学化. 然后用张量方法先对数据张量进行分解(希望找到一些潜在特征, 潜在关联), 再对分解后的进行SVM分类, 取得了比较好的分类效果. 这些效果的ground truth是通过问卷人工标定的. 此外, 对于他们定义的特征, 也进行了大量实用性分析. 应该是经过了漫长的科研历程才能做出的结果.
- Computers & Education 2009 *Exploring user experiences as predictors of MMORPG addiction* 社会学文章, 有一系列的, 不多加赘述, 套路类似. 自己通过一些调研, 找出可能的影响的predictors, 再用简单的回归去验证这些特征是否有效. 如果我们要像上面这篇文章这么做的话, 可能还是要收集些类似的文章, 相当于是做特征定义. 在曹楠老师的两篇异常检测中都有大量的特征定义.

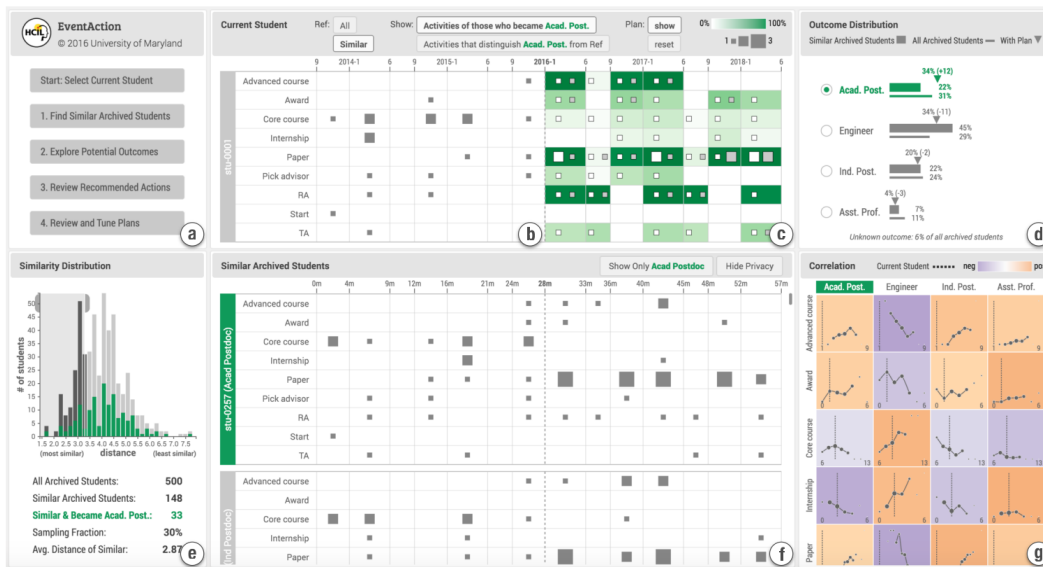


图 1: EventAction

- VAST2016 *EventAction: Visual Analytics for Temporal Event Sequence Recommendation* 首先看系统做的是很美观的(如上图1), 主题做时间序列推荐也是很新颖的, 文章写的也不错内容饱满, 每个图例注释都很优雅. 大问题是时序推荐, 文中叙述时候都是在讲一个实例, 类似于学生生涯规划, 连workflow 也是基于此. 对于时间序列本身的表示, 他们的关联, 提供的交互优点很多. 如果说为什么不能上tvcg我估计可能与描述问题实在太小, 用的度量太简单, 然后时间序列表示可能不够紧凑有关.
- VINCI2016 *Visual Analysis of User-Driven Association Rule Mining* 文章是对关联规则一些可以加入人主观思考的地方进行可视化的可视分析工作. 文章描述通俗易懂, 并对基础关联规则上进行了很多补充, 这些补充在系统中一一有所体现. 文章提到, 对于基础性的关联规则挖掘方法具有普适性, 这一点很有大局观, 也避免了reviewer的驳斥. 基于上下文的分布视图, 以及基于补充内容的交互是文章亮点. 基于矩阵的方格图最近似乎也常见, 毕竟很直观, 在空间足够时候是非常好用的. 当然关联规则套路很深, 对于supp和conf, 在一些进阶的方法中是非常复杂的, 这时候这个系统可能没那么有效了.
- VAST2016/TVCG *ViDX: Visual Diagnostics of Assembly Line Performance in Smart Factories* 智慧工厂, 工业流水线的可视诊断(也是找异常), 可能算是比较新的问题. 系统主要部分是基于Marey Graph (以前没见过) 的, 很直接虽然算不上特别美观, 但表现力也足够. 周围的小视图提供了足够的信息. 提供了两种不错的用于异常探测的交互技术quantile brush与sample brush, 基于两种算法, 具有一定普适性但作者自己也提到可能处理不了量特大的数据.